

09/554945  
528 Rec'd PCT/PTO 22 MAY 2000

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

- (i) APPLICANT: SHANGHAI SECOND MEDICAL UNIVERSITY
- (ii) TITLE OF THE INVENTION: A HUMAN HSG.III GENE
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Ratner & Prestia
  - (B) STREET: P.O. Box 980
  - (C) CITY: Valley Forge
  - (D) STATE: PA
  - (E) COUNTRY: USA
  - (F) ZIP: 19482
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Diskette
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: DOS
  - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: TO BE ASSIGNED
  - (B) FILING DATE:
  - (C) CLASSIFICATION: UNKNOWN
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER:
  - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Prestia, Paul F
  - (B) REGISTRATION NUMBER: 23,031
  - (C) REFERENCE/DOCKET NUMBER: GP-70526
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 610-407-0700
  - (B) TELEFAX: 610-407-0700
  - (C) TELEX: 846169

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2017 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

002250-54645560

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

TAAAGCTACG CCCTGGCCGC AGTCTCCGCG TCACAGGAAC TTCAGCACCC ACAGGGCGGA      60
CAGCGCTCCC CTCTACCTGG AGACTTGACT CCCGCGCGCC CCAACCCTGC TTATCCCTTG      120
ACCGTCGAGT GTCAGAGATC CTGCAGCCGC CCAGTCCCAG CCCCTCTCCC GCCCCACACC      180
CACCCCTCCTG GCTCTTCTCG TTTTACTCC TCCTTTTCAT TCATAACAAA AGCTACAGCT      240
CCAGGAGCCC AGCGCCGGGC TGTGACCCAA GCCGAGCGTG GAAGAATGGG GTTCTCTGGG      300
ACCGGCACTT GGATTCTGGT GTTAGTGCTC CCGATTCAAG CTTTCCCCAA ACCTGGAGGA      360
AGCCAAGACA AATCTCTACA TAATAGAGAA TTAAGTGAG AAAGACCTTT GAATGAACAG      420
ATTGCTGAAG CAGAAGAAGA CAAGATTAAA AAAACATATC CTCCAGAAAA CAAGCCAGGT      480
CAGAGCAACT ATTCTTTTGT TGATAACTTG AACCTGCTAA GGGCAATAAC AGAAAAGGAA      540
AAAATTGAGA AAGAAAGACA ATCTATAAGA AGCTCCCCAC TTGATAATAA GTTGAATGTG      600
GAAGATGTTG ATTCAACCAA GAATCGAAAA CTGATCGATG ATTATGACTC TACTAAGAGT      660
GGATTGGATC ATAAATTTCA AGATGATCCA GATGGTCTTC ATCAACTAGA CGGGACTCCT      720
TTAACCGCTG AAGACATTGT CCATAAAATC GCTGCCAGGA TTTATGAAGA AAATGACAGA      780
GCCGTGTTTG ACAAGATTGT TTCTAAACTA CTTAATCTCG GCCTTATCAC AGAAAGCCAA      840
GCACATACAC TGGGAAGATGA AGTAGCAGAG GTTTTACAAA AATTAATCTC AAAGGAAGCC      900
AACAAATTATG AGGAGGATCC CAATAAGCCC ACAAGCTGGA CTGAGAATCA GGCTGGAAAA      960
ATACCAGAGA AAGTGACTCC AATGGCAGCA ATTCAAGATG GTCTTGCTAA GGGAGAAAAAC     1020
GATGAAACAG TATCTAACAC ATTAACCTTG ACAAATGGCT TGGAAAGGAG AACTAAAACC     1080
TACAGTGAAG ACAACTTTAG GGAATTCCAA TATTTCCCAA ATTTCTATGC GCTACTGAAA     1140
AGTATTGATT CAGAAAAAGA AGCAAAAGAG AAAGAAACAC TGATTACTAT CATGAAAACA     1200
CTGATTGACT TTGTGAAGAT GATGGTGAAA TATGGAACAA TATCTCCAGA AGAAGGTGTT     1260
TCCTACCTTG AAAACTTGGA TGAATGATT GCTCTTCAGA CAAAAACAA GCTAGAAAAA     1320
AATGCTACTG ACAATATAAG CAAGCTTTTC CCAGCACCAT CAGAGAAGAG TCATGAAGAA     1380
ACAGACAGTA CCAAGGAAGA AGCAGCTAAG ATGGAAAAGG AATATGGAAG CTTGAAGGAT     1440
TCCACAAAAG ATGATAACTC CAACCAGGA GGAAGACAG ATGAACCCAA AGGAAAAACA     1500
GAAGCCTATT TGGGAAGCCAT CAGAAAAAAT ATTGAATGGT TGAAGAAACA TGACAAAAAG     1560
GGAAATAAAG AAGATTATGA CCTTTCAAAG ATGAGAGACT TCATCAATAA ACAAGCTGAT     1620
GCTTATGTGG AGAAAGGCAT CCTTGACAAG GAAGAAGCCG AGGCCATCAA GCGCATTAT      1680
AGCAGCCTGT AAAAAATGCA AAAGATCCAG GAGTCTTTCA ACTGTTTCAG AAAACATPAT      1740
ATAGCTTAAA ACACCTCTAA TTCTGTGATT AAAATTTTTT GACCCAGGG TTATTAGAAA     1800
GTGCTGAATT TACAGTAGTT AACCTTTTAC AAGTGGTTAA AACATAGCTT TCTTCCCCTA     1860
AAAACATCTT GAAAGTAAAG TTGTATGTAA GCTGAGATTT TGTATACAGG AATCCTTATT     1920
TCCTCATAGN CTTATTATTT TATAATCAGG AATATGTTGC TTTGGAAAAA GCCTCTTAAT     1980
GGGCTGACCN TAAAACTCA ATCNCCTTC CACTGTC      2017

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 468 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Gly Phe Leu Gly Thr Gly Thr Trp Ile Leu Val Leu Val Leu Pro
 1           5           10           15
Ile Gln Ala Phe Pro Lys Pro Gly Gly Ser Gln Asp Lys Ser Leu His
          20           25           30
Asn Arg Glu Leu Ser Ala Glu Arg Pro Leu Asn Glu Gln Ile Ala Glu

```

3/3

## SEQUENCE INFORMATION

SEQ ID NO:1

5 TAAAGCTACGCCCTGGCCGCGAGTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGA  
CAGCGCTCCCCCTCTACCTGGAGACTTGACTCCCCGCGCGCCCAACCTGCTTATCCCTTC  
ACCGTTCGAGTGTCTAGAGATCCTGCAGCCGCCAGTCCCGGCCCTCTCCCGCCCCACACC  
CACCCCTCCTGGCTCTTCTGTGTTTTACTCCTCCTTTTCATTTCATAACAAAAGCTACAGCT  
CCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGGAAGAATGGGGTTCTCTGGG  
ACCGGCACTTGGATTCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCAACCTGGAGGA  
AGCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAG  
10 ATTGCTGAAGCAGAAGAAGACAAGATTAACAAAACATATCCTCCAGAAAACAAGCCAGGT  
CAGAGCAACTATTCTTTTGTGATAACTTGAACCTGCTAAGGGCAATAACAGAAAAGGAA  
AAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCACTTGATAATAAGTTGAATGTG  
GAAGATGTTGATTCAACCAAGAATCGAAAAGTATGATGATTATGACTCTACTAAGAGT  
GGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATCACTAGACGGGACTCCT  
15 TTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAGAAATGACAGA  
GCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGCCAA  
GCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCC  
AACAATTATGAGGAGGATCCCAATAAGCCACAAAGCTGGACTGAGAATCAGGCTGGAAAA  
ATACCAGAGAAAGTGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAA  
20 GATGAAACAGTATCTAACACATTAACCTTGACAAATGGCTTGGAAGGAGAACTAAACCC  
TACAGTGAAGACAACTTTAGGGACTTCCAATATTTCCCAATTTCTATGCGCTACTGAAA  
AGTATTGATTGAGAAAAGAAGCAAAAGAGAAAGAAACCTGATTACTATCATGAAACA  
CTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCTCCAGAAGAAGGTGTT  
TCCTACCTTGAAAAGCTTGATGAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAA  
25 AATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAA  
ACAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGAT  
TCCACAAAAGATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACA  
GAAGCCTATTTGGAAGCCATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAG  
GGAAATAAAGAAGATTATGACCTTTCAAAGATGAGAGACTTCATCAATAACAAGCTGAT  
30 GCTTATGTGGAGAAAGGCATCCTTGACAAGGAAGAAGCCGAGGCCATCAAGCGCATTTAT  
AGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAACTGTTTCAGAAAACATAAT  
ATAGCTTAAACACTTCTAATTCTGTGATTAAATTTTTTGACCCAAGGGTTATTAGAAA  
GTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTCCCGTA  
AAAATATCTGAAGGTAAAGTTGTATGTAAGCTGAGATTTTGTATACAGGAATCCTTATT  
35 TCCTCATAGNCTTATTATTTTATAATCAGGAATATGTTGCTTTGGAAGAAAGCCTCTTAAT  
GGGCTGACNTAAAACTCAATCCNTCTTCCACTGTC

SEQ ID NO:2

MGFLGTGTWILVVLVLP IQAFFPKPGGSODKSLHNRELSAERPLNEQIAEAEEDKIKKTYFP  
ENKPGQSNYSFVDNLNLLRAITEKEKIEKERQSISSPLDNKLNVEDVDSTKNRKLIDDY  
DSTKSGLDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL  
5 ITESQAHTLEDEVAEVLQKLISKEANNYBEDPNKPTSWTENQAGKIEKVTPMAAIQDGL  
AKGENDETVSNTLTLTNGLERRTKTYSIEDNFRDPQYFPNFYALLKSIDSEKEAKEKETLI  
TIMKTLIDFVKMMVKYGTISPEEGVSYLENLDEMIALQTKNLEKNATDNISKLEFPAPSE  
KSHEETDSTKEEAAMKEYGSLKDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLK  
KHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL